UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2008/2009

November 2008

MSG 366 – Multivariate Analysis
[Analisis Multivariat]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of TWENTY TWO pages of printed material before you begin the examination.

[Sila pastikan bahawa kertas peperiksaan ini mengandungi DUA PULUH DUA muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

Instructions: Answer **all four** [4] questions.

*[**Arahan:** Jawab **semua empat** [4] soalan.]*

1. (a) Let $\mathbf{X}' = (X_1, X_2, X_3)$ is distributed as $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where

$$\boldsymbol{\mu} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 4 \end{bmatrix}.$$

(i) Find the distribution of $5X_1 - 4X_2 + 3X_3$.

(ii) Find the correlation between X_1 and $\frac{1}{2}X_2 + \frac{1}{2}X_3$.

[30 marks]

(b) Let $\mathbf{X}$ be $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu}' = [5 \quad 5 \quad 10 \quad 10]$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 10 & 4 & 4 & 4 \\ 4 & 20 & 10 & 10 \\ 4 & 10 & 20 & 10 \\ 4 & 10 & 10 & 10 \end{bmatrix}.$$

Determine whether each of the following random variables are independent. Explain.

(i) X_1 and X_3.

(ii) X_2 and X_4.

(iii) (X_1, X_3) and (X_2, X_4).

(iv) X_1 and $X_1 + 3X_2 - 2X_3 + X_4$.

[40 marks]

(c) Given the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

(i) Is $\mathbf{A}$ symmetric?

(ii) Is $\mathbf{A}$ positive definite?

(iii) Determine the spectral decomposition of $\mathbf{A}$.

[30 marks]

2. (a) Let $\mathbf{x}' = [4 \quad 1 \quad -2]$ and $\mathbf{y}' = [-1 \quad 3 \quad 1]$.

(i) Find the angle between $\mathbf{x}$ and $\mathbf{y}$.

(ii) Find the projection of $\mathbf{y}$ on $\mathbf{x}$.

[30 marks]

(b) Let $\mathbf{X}_1, \ldots, \mathbf{X}_{60}$ be a random sample of size 60 from an $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ population. Specify the distribution of the following:

(i) $(\mathbf{X}_1 - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_1 - \boldsymbol{\mu})$.

(ii) $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

(iii) $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

[30 marks]

1. (a) *Katakan* $\mathbf{X}' = (X_1, X_2, X_3)$ *adalah tertabur sebagai* $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *dengan*

$$\boldsymbol{\mu} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} \text{ dan } \boldsymbol{\Sigma} = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 4 \end{bmatrix}.$$

(i) *Cari taburan bagi* $5X_1 - 4X_2 + 3X_3$.

(ii) *Cari korelasi antara* X_1 *dan* $\frac{1}{2}X_2 + \frac{1}{2}X_3$.

[30 markah]

(b) *Katakan* $\mathbf{X}$ *adalah* $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *dengan* $\boldsymbol{\mu}' = [5 \quad 5 \quad 10 \quad 10]$ *dan*

$$\boldsymbol{\Sigma} = \begin{bmatrix} 10 & 4 & 4 & 4 \\ 4 & 20 & 10 & 10 \\ 4 & 10 & 20 & 10 \\ 4 & 10 & 10 & 10 \end{bmatrix}.$$

Tentukan samada pemboleh ubah berikut adalah tak bersandar. Jelaskan jawapan anda.

(i) X_1 *dan* X_3.

(ii) X_2 *dan* X_4.

(iii) (X_1, X_3) *dan* (X_2, X_4).

(iv) X_1 *dan* $X_1 + 3X_2 - 2X_3 + X_4$.

[40 markah]

(c) *Diberi suatu matriks* $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

(i) *Adakah* $\mathbf{A}$ *simetri?*

(ii) *Adakah* $\mathbf{A}$ *tentu positif?*

(iii) *Tentukan penghuraian spektrum bagi* $\mathbf{A}$.

[30 markah]

2. (a) *Katakan* $\mathbf{x}' = [4 \quad 1 \quad -2]$ *dan* $\mathbf{y}' = [-1 \quad 3 \quad 1]$.

(i) *Cari sudut antara* $\mathbf{x}$ *dan* $\mathbf{y}$.

(ii) *Cari unjuran bagi* $\mathbf{y}$ *terhadap* $\mathbf{x}$.

(b) *Katakan* $\mathbf{X}_1, \ldots, \mathbf{X}_{60}$ *adalah suatu sampel rawak bersaiz 60 dari suatu populasi* $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Tentukan taburan berikut:*

(i) $(\mathbf{X}_1 - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_1 - \boldsymbol{\mu})$.

(ii) $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

(iii) $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

[30 markah]

(c) Given the data matrix for a random sample of size $n = 3$ from a normal bivariate population:

$$\mathbf{X} = \begin{bmatrix} 12 & 17 \\ 18 & 20 \\ 14 & 16 \\ 20 & 18 \\ 16 & 19 \end{bmatrix}$$

(i) Evaluate Hoteling T^2 to test $H_0 : \mu' = [18 \quad 16]$

(ii) Specify the distribution of T^2 for the situation in (i). State your assumptions.

(iii) Using (i) and (ii), test H_0 at $\alpha = 0.05$. Give your conclusion.

[40 marks]

3. (a) Observations on two responses are collected for the following three treatments. The observation vectors $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ are

Treatment 1: [4 9] [7 9]
Treatment 2: [1 6] [2 3]
Treatment 3: [3 1] [2 3]

(i) Build a one-way MANOVA table for this data.

(ii) Evaluate Wilks' Lambda and test for treatment effects at $\alpha = 0.01$.

(iii) Repeat the test using the chi-square approximation with Bartlett's correction.

[50 marks]

(b) Rates of glucose metabolism were measured in four regions of the brains of $N = 12$ subjects provided by the data and summary statistics of MINITAB output in APPENDIX A.

Obtain the 95% simultaneous confidence intervals and give your conclusion for each of the following:

(i) region difference,

(ii) side difference,

(iii) interaction between region and side.

[30 marks]

(c) *Diberi suatu matriks data bagi suatu sampel rawak bersaiz* $n = 3$ *dari suatu populasi normal bivariat:*

$$\mathbf{X} = \begin{bmatrix} 12 & 17 \\ 18 & 20 \\ 14 & 16 \\ 20 & 18 \\ 16 & 19 \end{bmatrix}$$

(i) *Nilaikan Hoteling* T^2 *untuk menguji* $H_0 : \boldsymbol{\mu}' = [18 \quad 16]$

(ii) *Tentukan taburan bagi* T^2 *bagi situasi di bahagian (i). Nyatakan andaian-andaian anda.*

(iii) *Dengan menggunakan (i) dan (ii), uji* H_0 *pada* $\alpha = 0.05$. *Beri kesimpulan anda.*

[40 markah]

3. (a) *Cerapan terhadap dua respons sambutan dikumpul bagi dua rawatan berikut. Vektor cerapan* $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ *adalah*

Rawatan 1:	[4 9]	[7 9]
Rawatan 2:	[1 6]	[2 3]
Rawatan 3:	[3 1]	[2 3]

(i) *Binakan suatu jadual MANOVA satu-hala bagi data ini.*

(ii) *Nilaikan Wilks' Lambda dan uji kesan rawatan pada* $\alpha = 0.01$.

(iii) *Ulang ujian tersebut dengan menggunakan penghampiran khi-kuasa dua dengan pembetulan Bartlett.*

[50 markah]

(b) *Kadar metabolisma glucosa telah diukur dalam empat kawasan otak bagi* $N = 12$ *subjek menghasilkan data and output statistik ringkasan MINITAB seperti di LAMPIRAN A.*

Dapatkan selang keyakinan serentak 95% dan beri kesimpulan anda bagi setiap yang berikut:

(i) *perbezaan kawasan,*

(ii) *perbezaan sisi,*

(iii) *saling tindak antara kawasan dan sisi.*

[30 markah]

(c) Explain why the cluster analysis might be performed and how it differ from the discriminant analysis. Give an example.

[20 marks]

4. (a) The Achievement Test scores of random samples of 15 boys and 15 girls from a class in a school are shown in the table below. The test has six sections: Vocabulary (**V**), Word Recognition Skills (**WREC**), Reading Comprehension (**RC**), Mathematical Concepts (**MCON**), Mathematical Computation (**MCOMP**), and Mathematical Problem Solving (**MPS**).

	V	WREC	RC	MCON	MCOMP	MPS
Girls						
1	638	617	631	615	524	541
2	638	607	631	591	508	572
3	530	576	569	573	501	496
4	480	501	541	521	501	496
5	571	576	610	539	470	561
6	638	617	717	581	516	586
7	614	558	604	591	548	576
8	638	630	631	602	508	572
9	638	558	616	581	516	561
10	587	564	585	573	470	510
11	515	546	548	533	501	541
12	614	673	616	591	575	586
13	638	617	652	581	548	561
14	638	617	694	685	595	633
15	638	630	652	615	608	586
Boys						
1	638	630	604	660	501	541
2	614	617	668	660	575	572
3	614	558	594	602	470	533
4	571	617	652	591	501	633
5	587	607	604	573	556	586
6	539	546	548	615	565	586
7	508	583	555	591	532	572
8	614	673	589	685	548	633
9	587	570	562	551	508	633
10	614	607	565	602	501	561
11	587	590	631	632	524	605
12	614	598	616	573	532	605
13	614	617	616	660	532	605
14	638	607	594	558	540	633
15	508	570	531	533	516	541

The SPSS factor analysis program was used to fit the factor models. The output are as in APPENDIX B. Interpret the results and give your conclusion.

[50 marks]

(c) *Terangkan kenapa analisis kelompok dijalankan dan bagaimana ianya berbeza dengan analisis pembezalayanan. Berikan contoh.*

[20 markah]

4. (a) *Skor suatu Ujian Pencacapaian bagi sampel rawak 15 lelaki dan 15 perempuan dari suatu kelas di sebuah sekolah ditunjukkan dalam jadual berikut. Ujian terdiri dari enam bahagian: Vocabulary (**V**), Word Recognition Skills (**WREC**), Reading Comprehension (**RC**), Mathematical Concepts (**MCON**), Mathematical Computation (**MCOMP**), dan Mathematical Problem Solving (**MPS**).*

	V	*WREC*	*RC*	*MCON*	*MCOMP*	*MPS*
Girls						
1	*638*	*617*	*631*	*615*	*524*	*541*
2	*638*	*607*	*631*	*591*	*508*	*572*
3	*530*	*576*	*569*	*573*	*501*	*496*
4	*480*	*501*	*541*	*521*	*501*	*496*
5	*571*	*576*	*610*	*539*	*470*	*561*
6	*638*	*617*	*717*	*581*	*516*	*586*
7	*614*	*558*	*604*	*591*	*548*	*576*
8	*638*	*630*	*631*	*602*	*508*	*572*
9	*638*	*558*	*616*	*581*	*516*	*561*
10	*587*	*564*	*585*	*573*	*470*	*510*
11	*515*	*546*	*548*	*533*	*501*	*541*
12	*614*	*673*	*616*	*591*	*575*	*586*
13	*638*	*617*	*652*	*581*	*548*	*561*
14	*638*	*617*	*694*	*685*	*595*	*633*
15	*638*	*630*	*652*	*615*	*608*	*586*
Boys						
1	*638*	*630*	*604*	*660*	*501*	*541*
2	*614*	*617*	*668*	*660*	*575*	*572*
3	*614*	*558*	*594*	*602*	*470*	*533*
4	*571*	*617*	*652*	*591*	*501*	*633*
5	*587*	*607*	*604*	*573*	*556*	*586*
6	*539*	*546*	*548*	*615*	*565*	*586*
7	*508*	*583*	*555*	*591*	*532*	*572*
8	*614*	*673*	*589*	*685*	*548*	*633*
9	*587*	*570*	*562*	*551*	*508*	*633*
10	*614*	*607*	*565*	*602*	*501*	*561*
11	*587*	*590*	*631*	*632*	*524*	*605*
12	*614*	*598*	*616*	*573*	*532*	*605*
13	*614*	*617*	*616*	*660*	*532*	*605*
14	*638*	*607*	*594*	*558*	*540*	*633*
15	*508*	*570*	*531*	*533*	*516*	*541*

Program analisis faktor SPSS digunakan untuk suaikan model faktor. Outputnya adalah seperti di LAMPIRAN B. Interpret keputusan dan beri kesimpulan anda.

[50 markah]

(b) A certain graduate program requires that applicants sit for a diagnostic test which includes Verbal, Quantitative, and Analytical test scores, although many other kind of information enter into the admissions decision. The Verbal (**V**), Quantitative (**Q**), and Analytic (**A**) percentiles for $N_1 = 8$ successful and $N_2 = 28$ unsuccessful applications in one year are given in the following table:

Admitted			Not Admitted		
V	**Q**	**A**	**V**	**Q**	**A**
83	96	87	27	98	84
94	93	95	36	93	76
40	95	78	31	93	70
91	92	95	36	95	65
91	96	84	18	96	78
65	98	91	82	95	90
87	90	97	23	98	86
91	93	81	28	88	19
			31	93	33
			93	96	98
			28	96	67
			37	69	70
			78	83	97
			67	94	81
			58	96	47
			25	79	17
			74	97	64
			48	98	98
			80	95	95
			69	95	42
			46	93	71
			18	94	72
			9	65	16
			95	83	99
			34	98	80
			73	88	67
			33	92	81
			69	97	81

The discriminant analysis is performed on this set of data. The generated SPSS output are as in APPENDIX C. Interpret the result and give your conclusion.

[50 marks]

(b) *Suatu program ijazah memerlukan pemohon menduduki suatu ujian daignostik yang mengandungi Verbal, Quantitative, dan skor ujian Analytical, walaupun berbagai maklumat lain diambil kira untuk keputusan penerimaan. Persentil bagi Verbal (V), Quantitative (Q), dan Analytic (A) percentiles bagi permohonan berjaya* $N_1 = 8$ *dan tidak berjaya* $N_2 = 28$ *dalam satu tahun diberi dalam jadual berikut:*

Diterima			***Tak Diterima***		
V	***Q***	***A***	***V***	***Q***	***A***
83	96	87	27	98	84
94	93	95	36	93	76
40	95	78	31	93	70
91	92	95	36	95	65
91	96	84	18	96	78
65	98	91	82	95	90
87	90	97	23	98	86
91	93	81	28	88	19
			31	93	33
			93	96	98
			28	96	67
			37	69	70
			78	83	97
			67	94	81
			58	96	47
			25	79	17
			74	97	64
			48	98	98
			80	95	95
			69	95	42
			46	93	71
			18	94	72
			9	65	16
			95	83	99
			34	98	80
			73	88	67
			33	92	81
			69	97	81

Analisis pembezalayanan dijalankan ke atas set data ini. Output SPSS yang dijana adalah seperti di LAMPIRAN C. Interpret keputusan dan beri kesimpulan anda.

[50 markah]

APPENDIX A / LAMPIRAN A

Row	Infe_Left	Infe_Right	Supe_Left	Supe_Right
1	3.06	3.15	4.35	4.33
2	2.69	3.04	3.74	3.66
3	3.83	4.10	5.33	5.60
4	4.31	4.32	5.63	5.93
5	2.45	2.72	3.46	3.75
6	3.50	3.84	5.45	5.76
7	3.53	3.62	5.13	5.35
8	5.28	5.35	6.47	6.63
9	3.28	3.46	4.62	4.65
10	2.58	2.69	3.85	3.97
11	2.28	2.20	3.49	3.62
12	1.89	1.79	2.60	2.86

Descriptive Statistics: Infe_Left, Infe_Right, Supe_Left, Supe_Right

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Infe_Lef	12	3.223	3.170	3.151	0.951	0.274
Infe_Rig	12	3.357	3.305	3.314	0.979	0.283
Supe_Lef	12	4.510	4.485	4.505	1.124	0.324
Supe_Rig	12	4.676	4.490	4.662	1.160	0.335

Covariances: Infe_Left, Infe_Right, Supe_Left, Supe_Right

	Infe_Lef	Infe_Rig	Supe_Lef	Supe_Rig
Infe_Lef	0.903933			
Infe_Rig	0.920148	0.959333		
Supe_Lef	1.034936	1.071082	1.263055	
Supe_Rig	1.057570	1.091685	1.295236	1.345027

Matrix COVA1

0.90393	0.92015	1.03494	1.05757
0.92015	0.95933	1.07108	1.09168
1.03494	1.07108	1.26305	1.29524
1.05757	1.09168	1.29524	1.34503

Matrix INVERSECOVA1

49.2669	-41.7748	0.8775	-5.6764
-41.7748	56.0907	-26.7143	13.0463
0.8775	-26.7143	96.0779	-71.5288
-5.6764	13.0463	-71.5288	63.4987

...11/-

APPENDIX B / LAMPIRAN B

Correlation Matrix

		V	WREC	RC	MCON	MCOMP	MPS
Correlation	V	1.000	.636	.712	.514	.270	.395
	WREC	.636	1.000	.549	.596	.439	.501
	RC	.712	.549	1.000	.460	.358	.382
	MCON	.514	.596	.460	1.000	.462	.387
	MCOMP	.270	.439	.358	.462	1.000	.484
	MPS	.395	.501	.382	.387	.484	1.000
Sig. (1-tailed)	V		.000	.000	.002	.075	.015
	WREC	.000		.001	.000	.008	.002
	RC	.000	.001		.005	.026	.019
	MCON	.002	.000	.005		.005	.017
	MCOMP	.075	.008	.026	.005		.003
	MPS	.015	.002	.019	.017	.003	

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.811
Bartlett's Test of Sphericity	Approx. Chi-Square	65.829
	df	15
	Sig.	.000

Anti-image Matrices

		V	WREC	RC	MCON	MCOMP	MPS
Anti-image Covariance	V	.387	-.135	-.232	-.078	.091	-.041
	WREC	-.135	.443	-.032	-.148	-.071	-.118
	RC	-.232	-.032	.458	-.020	-.086	-.019
	MCON	-.078	-.148	-.020	.563	-.156	-.004
	MCOMP	.091	-.071	-.086	-.156	.644	-.206
	MPS	-.041	-.118	-.019	-.004	-.206	.651
Anti-image Correlation	V	.746(a)	-.326	-.551	-.168	.181	-.082
	WREC	-.326	.850(a)	-.072	-.296	-.133	-.220
	RC	-.551	-.072	.793(a)	-.040	-.159	-.035
	MCON	-.168	-.296	-.040	.866(a)	-.259	-.006
	MCOMP	.181	-.133	-.159	-.259	.775(a)	-.319
	MPS	-.082	-.220	-.035	-.006	-.319	.856(a)

a Measures of Sampling Adequacy(MSA)

...12/-

Communalities

	Initial	Extraction
V	1.000	.850
WREC	1.000	.705
RC	1.000	.747
MCON	1.000	.582
MCOMP	1.000	.782
MPS	1.000	.636

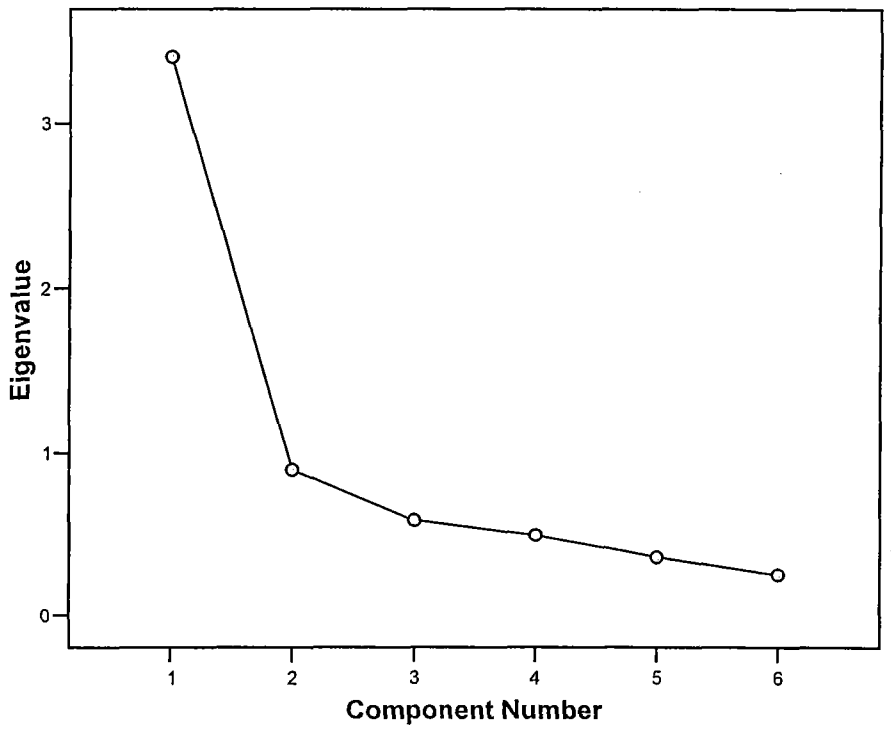
Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotatio
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	3.403	56.712	56.712	3.403	56.712	56.712	2.426
2	.899	14.975	71.687	.899	14.975	71.687	1.875
3	.593	9.876	81.564				
4	.498	8.300	89.864				
5	.358	5.968	95.831				
6	.250	4.169	100.000				

Extraction Method: Principal Component Analysis.

Scree Plot



Unrotated Component Matrix(a)

...13/-

	Component	
	1	2
V	.799	-.459
WREC	.838	-.062
RC	.778	-.376
MCON	.761	.051
MCOMP	.642	.607
MPS	.682	.413

Extraction Method: Principal Component Analysis.
a 2 components extracted.

Rotated Component Matrix(a)

	Component	
	1	2
V	.911	.141
WREC	.693	.475
RC	.842	.192
MCON	.562	.515
MCOMP	.122	.876
MPS	.274	.749

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 3 iterations.

Component Transformation Matrix

Component	1	2
1	.781	.625
2	-.625	.781

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Between-Subjects Factors

		N
Gender	1	15
	2	15

Descriptive Statistics

	Gender	Mean	Std. Deviation	N
Factor 1	1	.1792584	1.06456583	15
	2	-.1792584	.93234677	15
	Total	.0000000	1.00000000	30
Factor 2	1	-.3327604	1.06262422	15
	2	.3327604	.83963136	15
	Total	.0000000	1.00000000	30

...14/-

Box's Test of Equality of Covariance Matrices(a)

Box's M	3.025
F	.930
df1	3
df2	141120.000
Sig.	.425

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a Design: Intercept+Gender

Multivariate Tests(b)

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.000	.000(a)	2.000	27.000	1.000
	Wilks' Lambda	1.000	.000(a)	2.000	27.000	1.000
	Hotelling's Trace	.000	.000(a)	2.000	27.000	1.000
	Roy's Largest Root	.000	.000(a)	2.000	27.000	1.000
Gender	Pillai's Trace	.148	2.341(a)	2.000	27.000	.115
	Wilks' Lambda	.852	2.341(a)	2.000	27.000	.115
	Hotelling's Trace	.173	2.341(a)	2.000	27.000	.115
	Roy's Largest Root	.173	2.341(a)	2.000	27.000	.115

a Exact statistic
b Design: Intercept+Gender

Levene's Test of Equality of Error Variances(a)

	F	df1	df2	Sig.
Factor 1	.063	1	28	.804
Factor 2	.945	1	28	.339

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Gender

...15/-

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	REGR factor score 1 for analysis 1	.964(a)	1	.964	.963	.335
	REGR factor score 2 for analysis 1	3.322(b)	1	3.322	3.622	.067
Intercept	REGR factor score 1 for analysis 1	.000	1	.000	.000	1.000
	REGR factor score 2 for analysis 1	.000	1	.000	.000	1.000
Gender	REGR factor score 1 for analysis 1	.964	1	.964	.963	.335
	REGR factor score 2 for analysis 1	3.322	1	3.322	3.622	.067
Error	REGR factor score 1 for analysis 1	28.036	28	1.001		
	REGR factor score 2 for analysis 1	25.678	28	.917		
Total	REGR factor score 1 for analysis 1	29.000	30			
	REGR factor score 2 for analysis 1	29.000	30			
Corrected Total	Factor 1	29.000	29			
	Factor 2	29.000	29			

a R Squared = .033 (Adjusted R Squared = -.001)
b R Squared = .115 (Adjusted R Squared = .083)

Estimates

Dependent Variable	Gender	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Factor 1	1	.179	.258	-.350	.708
	2	-.179	.258	-.708	.350
Factor 2	1	-.333	.247	-.839	.174
	2	.333	.247	-.174	.839

...16/-

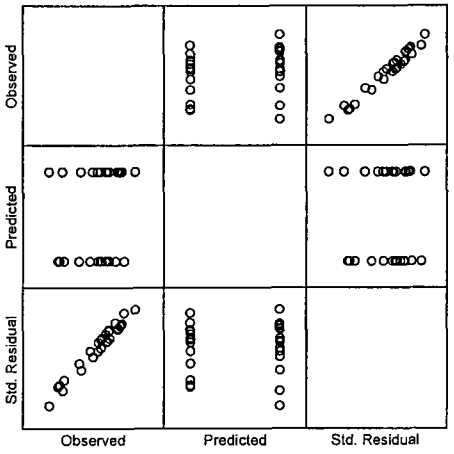
Pairwise Comparisons

Dependent Variable	(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	Sig.(a)	95% Confidence Interval for Difference(a)	
						Lower Bound	Upper Bour
Factor 1	1	2	.359	.365	.335	-.390	1.1
	2	1	-.359	.365	.335	-1.107	.3
Factor 2	1	2	-.666	.350	.067	-1.382	.0
	2	1	.666	.350	.067	-.051	1.3

Based on estimated marginal means

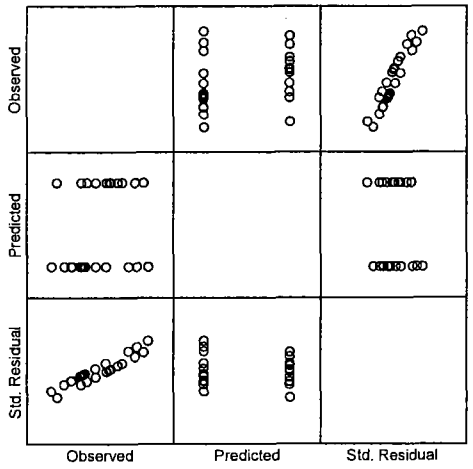
a Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Dependent Variable: REGR factor score 1 for analysis 1



Model: Intercept + Gender

Dependent Variable: REGR factor score 2 for analysis 1



Model: Intercept + Gender

...17/-

APPENDIX C / LAMPIRAN C

Group Statistics

Accept		Mean	Std. Deviation	Valid N (listwise) Unweighted	Weighted
0	V	48.07	25.156	28	28.000
	Q	91.32	8.420	28	28.000
	A	69.43	24.541	28	28.000
1	V	80.25	18.676	8	8.000
	Q	94.13	2.588	8	8.000
	A	88.50	7.091	8	8.000
Total	V	55.22	27.240	36	36.000
	Q	91.94	7.578	36	36.000
	A	73.67	23.223	36	36.000

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
V	.752	11.218	1	34	.002
Q	.976	.848	1	34	.364
A	.880	4.632	1	34	.039

Covariance Matrices(a)

Accept		V	Q	A
0	V	632.810	34.939	297.265
	Q	34.939	70.893	87.153
	A	297.265	87.153	602.254
1	V	348.786	-21.036	66.857
	Q	-21.036	6.696	-8.357
	A	66.857	-8.357	50.286
Total	V	742.006	38.784	351.790
	Q	38.784	57.425	75.067
	A	351.790	75.067	539.314

a The total covariance matrix has 35 degrees of freedom.

Box's Test of Equality of Covariance Matrices

Log Determinants

Accept	Rank	Log Determinant
0	3	16.650
1	3	11.073
Pooled within-groups	3	16.166

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

...18/-

Test Results

Box's M		22.584
F	Approx.	3.127
	df1	6
	df2	965.983
	Sig.	.005

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.344(a)	100.0	100.0	.506

a First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.744	9.617	3	.022

Standardized Canonical Discriminant Function Coefficients

	Function
	1
V	.884
Q	.084
A	.178

Structure Matrix

	Function
	1
V	.979
A	.629
Q	.269

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions Variables ordered by absolute size of correlation within function.

...19/-

Canonical Discriminant Function Coefficients

	Function
	1
V	.037
Q	.011
A	.008
(Constant)	-3.644

Unstandardized coefficients

Functions at Group Centroids

	Function
Accept	1
0	-.305
1	1.067

Unstandardized canonical discriminant functions evaluated at group means

Classification Function Coefficients

	Accept	
	0	1
V	.071	.121
Q	1.706	1.722
A	-.130	-.119
(Constant)	-75.805	-81.327

Fisher's linear discriminant functions

Classification Results(b,c)

			Predicted Group Membership		
		Accept	0	1	Total
Original	Count	0	19	9	28
		1	1	7	8
	%	0	67.9	32.1	100.0
		1	12.5	87.5	100.0
Cross-validated(a)	Count	0	18	10	28
		1	1	7	8
	%	0	64.3	35.7	100.0
		1	12.5	87.5	100.0

a Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
b 72.2% of original grouped cases correctly classified.
c 69.4% of cross-validated grouped cases correctly classified.

...20/-

APPENDIX D

The notations are as given in the lectures.

1. Spectral decomposition for a $k \times k$ symmetric matrix, A, is given by

$$A = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \ldots + \lambda_k e_k e_k'$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are eigenvalues of A and $e_1, e_2, \ldots, e_k$ are the corresponding standardized eigenvectors.

2. Suppose X has $E(X) = \mu$ and $\text{Cov}(\mathbf{X}) = \Sigma$. Thus $c'X$ has mean, $c'\mu$, and variance, $c'\Sigma c$.

3. Bivariate normal pdf:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}\left(1-\rho_{12}^2\right)}} \times \exp\left\{-\frac{1}{2\left(1-\rho_{12}^2\right)}\right.$$

$$\left.\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right) - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\}$$

4. Multivariate normal pdf:

$$f(x_1, x_2) = \frac{1}{(2\pi)^{p/2}\left|\sum\right|^{1/2}} \; e^{-(1/2)(x-\mu)'\sum^{-1}(x-\mu)}$$

5. If $X \sim N_p(\mu, \Sigma)$, then $AX \sim N_q(A\mu, A\Sigma A')$.

6. One-sampel :

(a) $$T^2 = n\left(\bar{X}-\mu\right)' S^{-1}\left(\bar{X}-\mu\right)$$

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j, \quad S = \frac{1}{n-1}\sum_{j-1}^{n}\left(X_j-\bar{X}\right)\left(X_j-\bar{X}\right)'$$

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

(b) Wilks' Lambda $\Lambda^{\frac{2}{n}} = \dfrac{\left|\hat{\Sigma}\right|}{\left|\hat{\Sigma}_o\right|} = \left(1+\dfrac{T^2}{n-1}\right)^{-1}$

(c) $100(1-\alpha)\%$ simultaneous confidence intervals for $\mathbf{a}'\boldsymbol{\mu}$:

$$\boldsymbol{a}'\bar{\boldsymbol{X}} \pm \sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)\;\; \boldsymbol{a}'\boldsymbol{S}\,\boldsymbol{a}}$$

(d) $100\,(1-\alpha)\%$ Bonferroni confidence interval for $\mu_i,\quad i=1,2,...,p$:

$$\bar{X}_i \pm t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{ii}}{n}}$$

(e) $100(1-\alpha)\%$ large sample confidence interval for $\mu_i : i=1,2,...p$

$$\bar{X}_i \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{ii}}{n}}$$

7. Paired comparisons

(a) $T^2 = n\left(\bar{\mathbf{D}}-\boldsymbol{\delta}\right)'\;\boldsymbol{\delta}_d^{-1}\left(\bar{\mathbf{D}}-\boldsymbol{\delta}\right)$

$$\bar{\mathbf{D}}=\frac{1}{n}\sum_{j=1}^{n}D_j \qquad \mathbf{S}_D=\frac{1}{n-1}\sum_{j=1}^{n}\left(D_j-\bar{\mathbf{D}}\right)\left(D_j-\bar{\mathbf{D}}\right)'$$

$$T^2 \sim \left[\frac{(n-1)p}{(n-p)}\right]F_{p,n-p}$$

(b) $100\,(1-\alpha)\%$ simultaneous confidence interval for δ_i:

$$\bar{d}_i \pm \sqrt{\frac{(n-1)p}{(n-p)}F_{p,n-p(\alpha)}}\sqrt{\frac{S_{d_i}^2}{n}}$$

8. Repeated Measure Design

(a) Let $\boldsymbol{C}$ be a contrast matrix

$$T^2 = n\left(\mathbf{C}\,\bar{\mathbf{X}}'\right)\left(\mathbf{CSC}'\right)^{-1}\mathbf{C}\,\bar{\mathbf{X}}$$

$$T^2 \sim \frac{(n-1)(q-1)}{(n-q+1)}F_{q-1,n-q+1}{}^{(\alpha)}$$

(b) $100\,(1-\alpha)\%$ simultaneous confidence intevals for $\mathbf{C}'\boldsymbol{\mu}$:

$$\mathbf{C}'\,\bar{\mathbf{X}} \pm \sqrt{\frac{(n-1)(q-1)}{(n-q+1)}F_{q-1,n-q+1}(\alpha)}\sqrt{\frac{\mathbf{C}'\mathbf{SC}}{n}}$$

9. Two independent samples:

(a) $$T^2 = \left[\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)\right]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_p\right]^{-1} \left[\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)\right]$$

$$T^2 \sim \frac{(n_1 + n_2 - 2)p}{(n_1 + n_2 - p - 1)} F_{p, n_1 + n_2 - p - 1}$$

(b) $100\,(1-\alpha)\%$ simultaneous confidence intervals for $a'(\mu_1 - \mu_2)$:

$$a'(\bar{X}_1 - \bar{X}_2) \pm c \sqrt{a'\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_p a}$$

where $c^2 = \dfrac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} \; F_{p, n_1 + n_2 - p - 1}$

10. One-way MANOVA:

(a) $$B = \sum_{\ell=1}^{g} n_\ell (\bar{x}_\ell - \bar{x})(\bar{x}_\ell - \bar{x})'$$

$$W = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (x_{\ell j} - \bar{x}_\ell)(x_{\ell j} - \bar{x}_\ell)'$$

$$\Lambda^* = \frac{|W|}{|B + W|}$$

(b) Bartlett:

$$-(n-1) - \frac{(p+g)}{2} \ln \Lambda^*$$

$$= -(n-1) - \frac{(p+g)}{2} \ln\left(\frac{|W|}{|B+W|}\right) \sim \chi^2_{P(g-1)}$$

(c) $100\,(1-\alpha)\%$ simultaneous confidence intervals for $\tau_{ki} - \tau_{\ell i}$:

$$\bar{X}_{ki} - \bar{X}_{\ell i} \pm t_{n-g}\left(\frac{\alpha}{pg(g-1)}\right) \sqrt{\frac{w_{ii}}{n-g}\left(\frac{1}{n_k} + \frac{1}{n_\ell}\right)}$$

$i = 1, 2, \ldots, p, \quad \ell < k = 1, 2, \ldots, g$

- ooo O ooo -